

MARTHA WEST  
Associate Professor  
Department of Librarianship  
San Jose State University  
San Jose, California  
and  
BRETT BUTLER  
Director  
Butler Associates

## **Performance Measures in Automated Systems Management**

When first approached as a possible speaker on the subject of performance benefit measures for library automation, my immediate reaction was that there are no such measures. Considerable cogitation, consultation, and survey of the literature hasn't made me change my mind, but my coauthor has convinced me that there is a beginning to the development of such measures in the actual management of automated systems today. This, then, is the thrust of our paper.

The reality of the situation in most American libraries today is that automated services are neither so well established nor so stable that normal professional management methods suffice for control and evaluation of the application. That is, the cost accounting and budget reporting which are standard in most business (and some public agency) environments are not suited to the management of rapidly changing operations. By normal standards of library operations, most automated services are characterized by a rate of change—of procedures, methods, services delivered, and costs—which can only be described as radical in the organizational sense.

Based on the limitation of the title of this paper, we are therefore limiting

the consideration of these management techniques to the citation of some relevant works in the bibliography. In this paper, we consider the management techniques related to what has become widely known as "program management"—the techniques of most direct relevance.

Having defined our scope by fiat as program management, it is useful to identify briefly what generally available techniques are used outside the information community to control programs. Table 1 highlights the features of seven management program evaluation techniques. Swanson has done an excellent job reducing the intellectual content of these techniques to a set of comparable abbreviated statements.<sup>1</sup> The most striking result is the interchangeability of most steps of most of the techniques described.

The rather scanty literature on performance evaluation in libraries can be classified into three groups. The *global* tries to measure performance related in some way to user satisfaction or societal goals. The *supervisory* discusses specific techniques for measuring productivity or effort in very specific environments, such as the relative value of various copiers for catalog card reproduction. Automated library systems in general do not fit well into either group noted above. Rather, they may be considered as projects which can be subjected to *strategic* evaluation. In other words, the automated system per se is designed to fit some broader set of policy goals; it is not reasonable to fault it for not delivering a service outside its design scope. For example, a management goal in an academic library may be: (1) to provide catalog access more effectively at the location of the present card catalog, or (2) to provide catalog access on a distributed basis all over the campus. The automated system selected can reasonably be evaluated only within the scope of the specific goal selected.

It is, of course, true that many automated systems are designed and built without such explicit identification of policy goals, but the shortcomings in the resultant operation are not failures in the performance of the automated system.

From a practical point of view, the choice of a specific evaluation technique within the library should depend primarily on evaluation of which technique is presently held in best repute by the library's parent organization. The actual use of any of these systematized guidelines, or any management-based logical alternative, will result in the provision of sufficient information for program management—if the technique is well and thoughtfully applied. None of these general techniques will substitute for intelligent thought about the library's specific needs and problems.

We are making an initial primary assertion regarding the management of automated library services: for most libraries, the automated services which are going to be used over the next few years will change with enough rapidity that each application should be considered as a specific program rather than a

<i>System Analysis (SA)<sup>74</sup></i>	<i>Operations Research (OR)<sup>1</sup></i>	<i>Benefit-Cost Analysis (BCA)<sup>142, 207</sup></i>
<ol style="list-style-type: none"> <li>1. Examine and clarify objectives; define issues of concern and problems</li> <li>2. Determine alternative courses of action that have some chance of resolving the issues</li> <li>3. Establish good criteria for choosing among the alternatives</li> <li>4. Obtain data (quantitative where possible) on the economic costs, effectiveness or benefits, and risks of the alternatives</li> <li>5. Construct models capable of predicting the consequences that are likely to follow from each choice of alternatives</li> <li>6. Compare results of applying the model to the alternatives in terms of the consequences</li> <li>7. Using predictions obtained from the models and other relevant information needed for further comparison of the alternatives, derive conclusions and recommend a course of action</li> <li>8. Test the conclusions wherever possible</li> </ol>	<ol style="list-style-type: none"> <li>1. Find the problem to be solved (decision to be made)</li> <li>2. Determine the objectives to be accomplished (outcomes desired)</li> <li>3. Identify the alternative courses of action available and the possible outcomes of each alternative</li> <li>4. Derive a measure of performance for evaluating the alternatives with respect to the probabilities of achieving the desired objectives</li> <li>5. Obtain information on controllable variables and on aspects of the environment (noncontrollable variables) that can affect the outcomes of the alternatives</li> <li>6. Construct a model that yields the performance measure as a function of the relationship among the variables</li> <li>7. Obtain solutions of the model, i.e., find values of the controllable variables that produce the best performance for specified values of the uncontrolled variables</li> <li>8. Select and implement the test solution</li> <li>9. Test the implementation; modify as necessary</li> </ol>	<ol style="list-style-type: none"> <li>1. Estimate the demand for proposed goods or services</li> <li>2. Determine alternative production possibilities</li> <li>3. Identify legal, resource, and technical constraints on the choice of alternatives</li> <li>4. Select a benefit-cost criterion measure for choosing among the alternatives</li> <li>5. Obtain information on the resource requirements, expected revenues, other direct benefits, possible side effects, intangibles, and uncertainties for each feasible alternative</li> <li>6. Compare and rank the alternatives on their benefit/cost ratios</li> <li>7. Include information on intangible (unpriced) and unmeasurable (not quantifiable) gain and loss factor not incorporated in benefit cost computations for each alternative</li> </ol>

<sup>74</sup>FISHER, GENE H. *Cost Considerations in Systems Analysis*. American Elsevier Publishing Co., Inc. New York, N.Y., 1971. 325 pp.

<sup>1</sup>ACKOFF, RUSSELL L.; SASIENI, MAURICE W. *Fundamentals of Operations Research*. John Wiley & Sons, Inc., New York, N.Y., 1968. 451 pp.

<sup>142</sup>MCKEAN, ROLAND N. *Efficiency in Government Through Systems Analysis with Emphasis on Water Resources Development*. A Rand Corporation Research Study. John Wiley & Sons, Inc., New York, N.Y., 1958. 336 pp.

<sup>207</sup>SEWELL, W. R. D.; DAVIS, J.; SCOTT, A. D.; ROSS, D. W. *Guide to Benefit-Cost Analysis*. Queen's Printer and Controller of Stationery, Ottawa, Canada, 1965. 49 pp.

<sup>238</sup>U.S. CONGRESS, SENATE COMMITTEE ON GOVERNMENT OPERATIONS. SUBCOMMITTEE ON NATIONAL SECURITY AND INTERNATIONAL OPERATIONS. *Planning-Programming-Budgeting: Budget Bureau Guidelines of 1968*. 90th Congress, 2d Session. Government Printing Office, Washington, D.C., 1968. 19 pp.

<sup>92</sup>HELLER, EDWARD D. *Value Management: Value Engineering and Cost Reduction*. Addison-Wesley Publishing Co., Reading, Mass., 1971. 229 pp.

<sup>9</sup>APPLEY, LAWRENCE A. "Standards of Management Performance." In: *Business Policy. A Management Audit Approach*. Edited by William T. Greenwood. Macmillan Co., New York, N.Y., 1967. 55-58.

**Table 1. Evaluation Techniques**

Planning- Programming- Budgeting (PPB) <sup>238</sup>	Value Analysis/ Engineering (VA/E) <sup>22</sup>	Management Audit (MA) <sup>9, 136</sup>	Evaluation (EVAL) <sup>218, 221, 259</sup>
<ol style="list-style-type: none"> <li>1. Identify Major Program Issues (MPI's) (questions requiring budget decisions)</li> <li>2. Produce program structures that show (1) program categories disclosing major objectives or scales of operation. (2) program subcategories or groupings of program elements according to output similarities, and (3) program elements or activities that yield a discrete output or end-product or related outputs</li> <li>3. Derive measurable program objectives; determine alternative programs to meet the objectives; collect data and perform analyses to compare and assess the costs, sources of funds, and anticipated benefits of each alternative</li> <li>4. Present decisions on choices in Program Memoranda (PM's) that summarize the analytic bases for the choices</li> <li>5. Produce a Program and Financial Plan (PFP) that relates data on outputs, costs, and financing sources for each program by program element for the past, current, and budget years and four future years</li> </ol>	<ol style="list-style-type: none"> <li>1. Obtain information on customer needs, specifications and preferences about products and/or services</li> <li>2. Obtain information on resource (input), production, and distribution costs</li> <li>3. Establish, compare and assess alternatives</li> <li>4. Reach a decision</li> <li>5. Plan, develop, and implement the alternative(s) decided upon</li> <li>6. Assess implementation; modify if necessary</li> </ol>	<ol style="list-style-type: none"> <li>1. Obtain information on company policies, organization, operating methods, financial procedures, personnel practices, and physical facilities</li> <li>2. Obtain information on company's investors, suppliers, customers, and competitors and on company's industry</li> <li>3. Identify problems and policy issues; prioritize them from major to minor; obtain facts and reasons for their existence</li> <li>4. Analyze and appraise company's condition and industry position</li> <li>5. Establish, compare, and assess alternative solutions to problems and issues</li> <li>6. Prepare recommendations for revisions or change</li> </ol>	<ol style="list-style-type: none"> <li>1. Identify system to be studied, its elements their characteristics, and purposes of evaluation</li> <li>2. Specify criteria, criterion variables and decision rules; state assumptions on which these will be based</li> <li>3. Obtain information on system variables (input, output, transactional, and intervening variables that may influence system operations and performance) and on the antecedent system if applicable and possible</li> <li>4. Compare outcomes with criteria on the decision rules, draw conclusions, make recommendations</li> <li>5. Assemble analyzed data in a format appropriate for decision-makers; disseminate report</li> <li>6. Design and assess alternatives to produce needed modifications or changes</li> <li>7. Reach a decision</li> <li>8. Plan, develop, and implement alternative(s) decided upon</li> <li>9. Assess implementation; modify if necessary</li> </ol>

<sup>136</sup>MARTINDELL, JACKSON. "The Management Audit." In *Business Policy, A Management Audit Approach*. Edited by William T. Greenwood, Macmillan Co., New York, N.Y. 1967, 83-91.

<sup>218</sup>STAKE, ROBERT E. "The Countenance of Educational Evaluation." *Teachers College Record*, 68:7 (April 1967) 523-540.

<sup>221</sup>STUFFLEBEAM, DANIEL L.; FOLEY, WALTER J.; GEPHART, WILLIAM J.; GUBA, EGON G.; HAMMOND, ROBERT L.; MERRIMAN, HOWARD O.; PROVUS, MALCOLM M. *Educational Evaluation and Decision Making*. F.E. Peacock Publishers, Inc., Itasca, Ill., 1971, 367 pp.

<sup>259</sup>WEISS, CAROL H. *Evaluation Research, Methods for Assessing Program Effectiveness*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1972, 160 pp.

**Table 1. Evaluation Techniques**

Source: Swanson, Rowena W. "Design and Evaluation of Information Systems." *Annual Review of Information Science and Technology* 10:54-55, 1975.



permanent service or fixture of the library's operation. For example, we would apply normal cost analysis techniques to the measure of performance in card filing if we assumed the card catalog would be a fixture of the specific library under study. If we assume implementation of any alternative to the card catalog, however, we can expect to see the use of several automated services over the coming decade. In a recent interview with public librarians who had been managing a book-catalog system for several years, we were discussing acceptance of the newly installed microform catalog. One person commented: "The patrons used to ask for the card catalog; now they ask where the book catalog went." The card catalog had lasted a century; the book catalog disappeared after a decade. There is no indication in the technology that this library (or any other) should expect a diminishing rate of change.

If this assertion is valid, and if the presentations by the other speakers confirm the growing feasibility of expanding the library's use of automated services, then the management of benefit analysis becomes a program management function.

While the jargon of professional business management is at least as confusing as the language of librarianship, the English-language definition of *program* delineates our intent sufficiently (in fact, the three disparate elements illustrate facets of the topic): (1) a prospectus or syllabus, (2) the events or pieces...of an entertainment or ceremony, and (3) a plan of procedure. There is one aspect common to these three definitions: a program is finite but not (except in the past tense) completed. A program is a project or an undertaking rather than a function. Collection development is a function of the library; a review of the holdings in a specific field against a standard list is a project.

Neither the history of automated library programs, nor the projections of likely changes in automated services, suggest that these programs can become functions. Change will occur rapidly enough so that a specific automated project can expect to have a lifespan of no more than a few years.

There are a number of program management and evaluation techniques which are widely used (or at least discussed) within the data processing industry, where rapid and continual change is accepted as normal. As with general management tools, we are excluding review of these techniques here on three related grounds:

1. Only a very few library systems have or will have the responsibility (or luxury) of selecting and controlling the computer operations environment within which they will operate. It does little good to know a great deal about structured programming management, if you buy your services from a book catalog vendor or OCLC.

2. Program management techniques rarely provide specific methods appropriate to library applications per se particularly because library applications tend to be more complex in terms of textual and language processing requirements than typical data processing jobs.
3. In our opinion, technicians both inside and outside the library community have overstated the importance to library administrators of knowing the relative efficiency of various computer equipment and software elements.

The key elements in program management of automated library applications are, rather, involved with how the library management defines its goals in specifying a new (automated) service, and then measures the results. This process is performance measurement. The term *performance measurement* necessarily implies the management of the program: else why bother taking the measure?

We are really, then, only interested in two elements. We want to measure the potential of the automation product to be procured, to define its potential task in the organization, and to define its costs. We then want to measure the results, in two major ways: (1) the degree to which the product delivered meets the specification; and (2) the way in which the product meets the target goals.\*

Between goals and specifications lies a tricky gray area: one must communicate the initial design concept to others. This may involve communicating a technical systems concept up the organization chart to management or horizontally to other staff. It may involve explaining a management concept—balancing collection to measured user demand—to several different staff groups. In the areas of interest here, it almost certainly will involve communicating to “vendors,” whether private contractors or one’s own agency data processing operation.

In fact, it is probably the difficulties of keeping track of events through this gray area that accounts for the fact that so few projects report both prior planning and subsequent evaluations. We have a lot of literature discussing project plans, and there is a substantial amount of analysis of existing services. Whenever one organization has lived through the real implementation of an abstract plan, groping through the gray areas (where the fine black-and-white plans dissolve or at least become confused) seems to sap the organization’s energies below the level required to complete the follow-up performance evaluation.

This observation does not particularly suggest that implementing an automated system is generally such a traumatic and unsuccessful experience that nobody wants to talk about it afterward. It seems more to be the result of

---

\*It is not necessary for a product to meet specifications in order to meet performance goals; and conversely, a product built exactly to specification may fail to meet the organization’s goals.

two factors: (1) as noted above, automated library systems are projects; new projects arrive and absorb the time which otherwise might be allocated to evaluating the old; and (2) few automation projects specifically include the formal evaluation phase in the project development plan.

### **Program Development**

Evaluation and measurement of targets and results is probably more important during implementation than during planning or after operations are established. It is not difficult to envision a number of ways in which automated services can improve library operations, although the selection of the specific alternative can present an extremely difficult decision. In addition, the evaluation of results in the broader sense is limited by the ability of the library to implement parallel experimental designs, so operating evaluations are primarily useful for management purposes or for defining a new project.

If the proper evaluation of results being obtained during implementation is not made, however, one can experience failures in performance ranging from the catastrophic (never reaching the operating stage) to the merely annoying ("Why didn't we remember to include that data in the conversion?").

It is axiomatic that before a program can be evaluated it must be defined and understood. An automated library system must be defined both in terms of present operations and desired goals. It is nonetheless necessary to present this axiom because there are so many illustrations of automated applications where only a portion of the goals are explicitly identified and where, therefore, full evaluation cannot be made. While other papers at this clinic illustrate quite clearly the processes used to quantify the goal of "saving money," too often the other major goal of a system project is expressed at the level of "increased services," which is difficult, if not impossible, to measure.

For instance, in implementing a cataloging system, does the service goal relate to increasing the convenience of use for the cataloging department, for the public, or for some combination? To answer this type of question, it is necessary to define each program component into discrete units which can be compared in a manual vs. automated mode with regard to: (1) old functions partially or totally discontinued, (2) new functions established, (3) functions with little or no change, and (4) revised procedures within a function. For each of these, it is necessary to identify the overhead or fixed portions of the functions, so the analysis can be quite clear with regard to which functions are potentially variable. Otherwise we fall into the fallacy of what Allen Veaner has termed the "anyhow" school of economics, illustrated by the classic phrase, "But we have the computer anyhow, so it won't cost us anything to...."



Fundamental to this analysis is the identification, formulation, and communication to organization members of program goals as related to the present system and as projected. These goals also need to be related to the parent organization's goals. It is essential to make clear whether the automated system is merely changing procedural or administrative processes by which present goals are being met, or whether the system is implementing new goals. One indication is that planning described almost entirely in terms of relative costs generally implies continuation of present goals, with the major change in the automated system seen as cost control.

Early automated systems almost entirely represented continuation of general library organization goals. This is illustrated by early efforts to automate library catalogs primarily to eliminate filing costs. Current efforts are much more involved in fulfilling expanded service goals such as the effort to automate library catalogs in order to decentralize access. Because the automated system in the current environment cannot be compared directly to the old manual processes (the automated orange is different than the manual apple), it is particularly important to specify the degree and manner in which it is expected the library operation will change. This cannot be done after development and implementation have occurred.

It is likely that some of the desired goals will be extremely difficult to quantify. There are basically three areas to consider during program development: (1) the "hard" evaluation data available or obtainable, (2) the "soft" evaluation information, and (3) the definition of the local environment (technical and political). As available information is "hard" and "soft," so some performance measures will be "hard" or "soft," depending on the level of management confidence in the information obtained. But consideration of these three areas during program development will at least provide a framework for management.

Various definitions of management focus on an identification of the key management role as "decision-making in the presence of insufficient information," or a variation of the phrase. So it is the recognition of the range in which "hard" information becomes "soft" for evaluation purposes that is important. The effort to gain perfect information in order to make a perfect decision results in no decision being made at all.

### **Hard Evaluation Data**

In general, costs are thought of as "hard" data; the mere presence of numeric data implies quantification (to an often unwarranted degree). It is the basic cost information which is being increasingly reviewed by the library's "controllers" (boards of trustees, regents, county supervisors, or legislators) who are experiencing increased problems in stretching the various taxpayers' dollars. It is, for instance, unlikely that we will see many automated library systems implemented in the next decade which baldly



admit to an increased level of cost. At the very least, the most sophisticated descriptions now stress reduction in the rate of cost increase, or in the per-unit cost of operation. The latter formulation, first promulgated regularly during the development of the BALLOTS system, is also cited by OCLC and (with supporting real operating figures) by the Ohio State University circulation system.

In cost analysis, it is unfortunately necessary to know what our existing costs are, a process much less glamorous than budgeting new future programs. While a number of methodologies exist for determining costs (and the choice of which to use will lie within the expertise and constraints of the particular institution), a warning must be given. Because of the total absence of standardized task definitions within library operations (Have we been able to reach a consensus on what constitutes original cataloging?) the local costs obtained are only going to be applicable to the particular institution collecting them. They cannot be compared exactly to the costs calculated at similar institutions and, indeed, probably cannot even be validated to proper experimental standards.

The goal for most libraries, however, is to establish the level or magnitude of costs to be compared. It is not reasonable to concern the evaluation effort with determination of current costs to the penny or 1 percent level of accuracy for two reasons. First, the comparative costs of the projected automated system probably cannot be projected with that level of accuracy. Second, the costs of obtaining the last few degrees of accuracy will be greater than the potential savings from that accuracy. We have recently been fortunate in working with some extremely explicit large-scale personnel cost figures,<sup>2</sup> but we would not have proposed the cost-gathering effort (and expense) for the sole purpose of our single-application automated systems study alone.

The value of the current operations' cost data lies in the use of cost levels as a type of benchmark; current costs are a measure of hard data to be used in evaluating proposals for a new automated system, and in evaluating the system once installed and operational.

It is important to note that the process of data collection, or observations of operations, can result in changes in the processes observed (the "Hawthorne effect" and more subtle variations). For this reason it is important to obtain the benchmark cost data before any of the changes related to the new system are implemented or discussed with staff. The observer will change the process somewhat by observation, and this effect cannot be removed entirely, but it is much worse if the "observer" is actively involved in changing processes while studying their costs and procedures. As noted above, one element to analyze is the use of revised procedures within an existing function. One must resist the temptation to do the revisions before measuring the starting point. (Conversely, evaluations often as-

sign benefits or cost reductions to new automated systems which are in fact nothing more than revisions of procedures which could have been accomplished with the manual system. If the cataloging department saves money by eliminating underlining of subject headings with a red pen, the cost saving cannot be attributed to the installation of the OCLC terminal.)

The following major cost elements should be collected to provide benchmark information for the evaluation process:

1. *Workload or demand.* Estimates can be provided from historical data for each item processed through the system. Because of the relatively large number of items in libraries' files, and because of the long historical time period most files represent, sophisticated sampling should be employed to ensure that historical data and current practice are related.

The objective of this element is to determine long-range trends in growth patterns and to determine the patterns of work-flow or demand over shorter periods. The second step is of particular importance because of the wide variations in demand or work-flow which characterize much library activity.

2. *Definition of process.* Current procedures must be analyzed to determine what work steps are actually involved; these may be assumed to have a tenuous relation to written work procedures or job descriptions. At a minimum, event frequencies estimated without measurement of actual experience may be considerably in error. Average times for each event may be obtained by sampling, time-and-motion studies, or diary methods.

At this point it is well to mention a resource which can and should be used for many cost procedures: plagiarism. A wide variety of detailed library cost studies have been published. Those who work with them can cite limits in each, and it is easy to identify areas where more needs to be done, but the cumulative information available is very much greater than was available a decade ago. In an individual library environment, if a piece of outside cost data looks reasonable in terms of limited local experience, it may be useful. If a number of independent pieces all confirm the initial local effort, do not be distracted in your evaluation process by the conviction that "our library is different." Use what is available and husband your resources for the next two steps in development and evaluation: evaluating that "soft" data and reviewing your local external environment.

3. *Defining interactions.* Independent of the specific procedures currently being followed, certain systematic interactions are taking

place which define the task being studied. These consist of a set of human interactions with: (a) other staff, (b) manual resources, and (c) computer-based resources. The system design will force a change in this mix. It is important to understand its present cost structure in order to predict these changes.

4. *Determining manpower requirements.* Tangible cost benefits from the implementation of an automated system have been most readily found through resulting changes in the mix of manpower skills. That is, while it is seldom realistic to project staff reductions, replacement of expensive staff by less expensive staff can be obtained for many functions. Therefore, it is necessary to define work rates and work standards for tasks through the methods cited above.
5. *Resource costs.* To staff costs must be added measures for unit hardware, software, and development costs for current activities. This consideration should include alternative costs for operation of current functions, i.e. leasing vs. buying equipment, etc.

Each of the five areas outlined should be examined in such a way that comparable information can be obtained about the present operating costs and projected systems costs. Otherwise we may have a four-way mix of automated oranges, manual apples, manual oranges, and automated apples—a fruit salad of analysis.

It is important to establish which “hard data” (as illustrated by costs in this discussion) properly belong in what Learman has called “the domain of analysis,” or “the sum of all valid areas of measurement.”<sup>3</sup> Learman outlines criteria for including information within a domain of analysis as:

- a) an identifiable, definable input and output for each application to be included; b) some proposed benefit which can be quantified for the proposed change; and c) some functional relationship or equivalence between an application in the current system and in the new system.<sup>4</sup>

Information which does not meet these criteria are outside Learman’s “domain” and in general belong in our classification of “soft information.” Unfortunately, many of the things management wants to know about a potential automated system lie in the area of “soft” information for most libraries undertaking current projects.

### **Soft Information**

There are two reasons that information is “soft” from the manager’s point of view: (1) either the questions being asked have not been defined fully, or (2) the resources do not exist to provide hard data on properly defined questions. For example, “How satisfied is the user of the library?”



A "hard" answer cannot be provided even with unlimited research funds until an agreed-upon definition of satisfaction is provided. An example of the second reason is the inability of most libraries to provide classic time-and-motion analysis on repetitive library tasks in order to obtain very precise operating marginal labor costs. We will discuss two areas of "soft" information which may be considered at the stage of program development—the evaluation of related library activities in other library systems and the evaluation of user needs, demands, or desires.

It is natural for a library to want to draw on the experience of other systems when planning a major change such as represented by almost any automated system. The insistence by some staff on the uniqueness of the local library is not often shared by the library administrator wishing to minimize risk and maximize results in an automated system. Those who are implementing pioneering systems soon become familiar with a wide range of visitors, national and international, who have arrived "to see how it is working at...."

Unfortunately, while library functions can be described as "the same" at some very broad level—everyone acquires, accounts for, catalogs, and circulates materials—these operations today differ widely due to the variety of institution goals they support and the level of instructional resources available. To obtain relatively "hard" information about, for instance, the performance and costs of a potential vendor's circulation system, would require that the library studied be similar in all features to the library looking for information.

In addition, it is increasingly true that the more successful automated systems are broader in scope and more complex than the earlier traditional automated functions. Many systems stretch the state of the technology, or at least the state of deliverable products. While it is true that those involved in the operation of such systems find it increasingly possible to evaluate and share each others' efforts (the transfer of the University of California, Berkeley, serials KWIC publication programs to Harvard is an example), such sharing does not suggest the evaluation process is simple for libraries without specialized internal systems staffs.

In our opinion, the reliance on other libraries' reported experiences, or on the results of "whirlwind" survey tours, are sufficiently limited to be a marginal source of reliable information to most libraries. Intelligent review of library and information technology literature can provide good indications of what is feasible, but a one-day site visit anywhere can provide only the politically useful "soft" information confirming that the site visited does in fact perform the functions the written literature claimed for it.

If libraries were organized as businesses or as mission-oriented federal agencies, they might have ongoing, patron-level, market research pro-



grams. Few libraries have any specific "hard" information beyond the demographic level, and no major library system in the United States yet admits to an ongoing program in this area. The efforts that do exist are concerned primarily with measuring patron "satisfaction" with a particular program (a film service, SDI awareness, etc.) or with patron reaction to a specific change in procedures such as the remote storage of little-used materials. This type of program evaluation is a limited effort which has little relationship to the general goals of the library, and thus is of limited value as input to management evaluation. Furthermore, it is often difficult to evaluate patron reaction to automated library services. The more basic measurements are obvious without consulting the patron; cutting in half the time required for charging a book can be measured readily in terms of staff savings, and some proportionate patron benefit can be assumed. As services move beyond housekeeping tasks, however, and into changes in patron access to the collection (new catalogs and automated reference services), evaluation of the effect of change becomes more difficult.

One problem with attempting to gain even "soft" information about patrons is the complexity of the behavioral science research which is required in order to obtain even minimally dependable answers. The number and variety of potential users are considerable (professional and clerical staff, casual and research patrons, young and old, minorities, students and businessmen, and the important majority group, nonusers). An added factor is the variation in use of the library over time (school years, seasons) which demands long-term measurement. The most complex problem, however, is that when dealing with planning for automated services, it is almost impossible to project accurately a potential service in a patron interview. Both patrons and the staff find it difficult to choose among equally unknown alternatives.

Where direct information can support the evaluation of a specific automated system proposal, it should be obtained. For instance, one could reasonably expect to obtain direct answers to the question, "How often do you use another library, and which one?" appropriately written out for presentation to patrons at a circulation desk. The answers could be quite relevant in evaluating the capability of an automated circulation system to maintain either (1) local-branch patron verification files, or (2) system-wide patron verification files. If a library's patrons used other branches in the system heavily, the latter option would be preferable.

Such limited evaluations of patron service will be replaced in most libraries by more abstract measures of services presently provided and to be included in a new automated service. For instance, Orr and his colleagues<sup>5</sup> have developed a general scheme of library services which includes a methodology for measuring collection, document delivery capability, and a

services inventory. It is practical and efficient for most library administrators to accept the basic outline of services provided by Orr (which includes providing documents, citations, answers, work space, and instruction, along with a miscellaneous class) and then to base their library's evaluation process on measures of these arbitrarily defined services. The alternative is for the local library to mount a public services behavioral research effort. The model for the library which does have some research capabilities and is not totally satisfied with the conclusions of others is available in the work of Ben-Ami Lipetz at Yale,<sup>6</sup> where a relatively limited scope of information was thoroughly documented from card catalog users. By keeping the scope of the research effort to a focus on "What did you want to find in the card catalog when you approached it?" rather than the much broader "What do you want in the library?" Lipetz produced perhaps the single most useful piece of research regarding user practice in the last decade.

### Definition of the Technical Environment

At this stage in the automated system's program, the library will have reviewed what can reasonably be found out about its present operations and the demands upon them, in terms of both "hard" data, such as costs, and "soft" data, such as patron market research. The most significant pitfalls, however, come in evaluating, from inside the organization, the technical resources available to the library to accomplish its goals. Various types of limits will be found to exist:

1. *Institutional*: the organization of the library within the parent organization; authority of the library over other related resources such as the computer center or library-like agencies within the parent organization; contractual or other ties to other libraries.
2. *Facilities*: availability of computer facilities; their type, configuration, and suitability for the peculiar data processing and transmission needs of library applications; policies regarding use of various computer languages; transmission links with other organizations, etc.
3. *Staff*: within the library and other related agencies in the parent organization, evaluation of the available staff's knowledge with regard to all phases of the required effort (systems analysis and design, programming, daily operations); staff ability to train others in the library and support user education; ability of the library to recruit, add, or train staff if required.
4. *File conversion*: relation of conversion for one application to future use for others; technical standards; use of in-house, vendor, or library network resources for conversion.
5. *Management priorities*: limits set by the parent organization or by

the library management which restrain certain applications (e.g., "all service shall be provided without charge"); availability of other resources which support cooperative development.

6. *Finances*: ability for financing capital investments in equipment, data conversion, and program start-up; sources for supporting operating expenses, including internal systems staff.

The end result of all this activity is to provide the management information necessary for the specification of an automation program. Evaluation of each step is determined by the adequacy of input for each of these areas: (1) decision criteria for internal vs. external development, (2) priorities for implementation of subprograms, (3) staffing projections, and (4) cost-benefit analysis. The process should result in the ability to use information gathered to write detailed program specifications, which can be used as a basis for "requests for proposals" and evaluation of vendor proposals.

### **Program Implementation**

The first step in implementation is often described as the "make-or-buy" decision; this is the choice between internal development of the automated system and contract outside development. In the practical sense, outside contractors should be defined to include the parent organization's data processing department, as well as independent vendors. The make-or-buy decision implies the evaluation of the performance capabilities of each type of resource (internal or contract) and in fact generally involves the evaluation of specific programs or services available from individual contractors. It must be recognized that for some library systems this step is functionally nonexistent or trivial. Local circumstances may include an administrative edict requiring all agencies within an organization to use the services of the organization's data processing department. In this case the library is mandated to buy from a specific vendor. Conversely, the first review of library and parent organization resources may demonstrate clearly that no necessary resources are available. In this case, independent sources are mandated.

The improper matching of internal and external resources has probably been responsible for more outright disasters in library automation applications than any other factor. When either good internal staff are burdened by poor computer access, a lack of competent staff leads to little control over vendor programs, or the library and the organization's data processing staff are competing for priorities, one can expect only minimally operational applications. Thus, it is prudent for the library's management to go through the processes of evaluating the implementation even if the choice of resources is forced administratively rather than occurring from exercise of library evaluations. It may be helpful the second time. If freedom of choice is available, the make-or-buy decisions can proceed using the library's choice



of the basic management techniques outlined in Table 1. If freedom of choice is not available, then "make" or do nothing.

In buying a contract service or product, there is a general distinction between standard and custom services. Real economies of scale are to be found in the design and delivery of standard services from computer systems to a large number of user libraries. Unfortunately, standardization of operations among most libraries has not progressed much beyond acceptance of the location for the rod hole in the catalog card. As a result, even the better-designed "standard" automated services tend to become "customized" over a period of time because of individual library customer needs or expressed desires.

Potential cost savings are one advantage of a standard contract service or product. A second benefit is that unless your library is the first customer, it is possible to determine that the described product is actually working. A third, and nontrivial, benefit is that the nature of the completed product is such that it can be fully specified, whereas a custom-designed product almost always suffers from some nonspecified discrepancies in expectations during development.

On balance, the evaluation of a system's performance is markedly easier if the system is a replication or transfer of a product or service installed elsewhere. The difficult point is to determine when seemingly innocent and trivial "local changes" transform the project from a standard to a custom system.

In evaluating make-or-buy choices, it will be necessary to evaluate closely the technological trends which define the technical environment of the proposed system. This paper has no room to discuss such important technical issues as microcomputers and value-added networks. It is perhaps worth observing that library automation efforts have more often "stretched" the state of computer technology than is expected by many in the computer or data processing profession. The development of extended character sets for computer printing and CRT display, the handling of very large files on-line for searching, and the use of commercial "light-pen" data entry devices are all examples from recent years. This situation places the library staff evaluating a project in the difficult situation of not wanting to design around obsolete equipment, but at the same time wanting to avoid the risks inherent in unproven technology. No simple method or formula will replace the use of knowledgeable specialists in analyzing the technology.

There are, however, some general guidelines which might be helpful in evaluating make-or-buy options:

1. *Fit your measures to your application.* The vendor procedures for hardware maintenance and error-correction are much more important in a circulation system than in a book catalog production service.



2. *Think of the next three projects.* If the short-term advantage of buying a standard circulation system is negated by the inability of that system to accept input from a later automated bibliographic system, the "standard" product may have saved your library little over making a custom product.
3. *Separate the application and the data base.* Over the past decade, probably the largest single waste of library automation efforts has come from the improperly evaluated development of a local library data base for a specific purpose, later found to be inappropriate for other uses or further implementation on different equipment. In evaluation, consider the development of the data base per se as a separate project from the initial application; but do force evaluation of the data base creation rather than just assuming that "then we can do anything."

### **Evaluation of Implementation**

Evaluating an automated project must be organized before the actual commitments to the program are made. If there is a single element of measuring the performance of a computer-based library application which is different from other data processing applications, it is the breadth of evaluation skills which may be required to assess the success of the project. This is partly because of practical problems in information handling and retrieval posed by large library files and processing demands, but it is also partly due to the complex and generally unevaluated set of cataloging procedures and administrative rules which have grown up in most libraries over a period of years.

It is redundant to insist that the first job of an evaluation effort should be to measure the performance of the present library system which is to be replaced. It is still common, however, to find situations where this practice has not been followed. For instance, during the first three years of on-line operations of the OCLC cataloging system, virtually no studies were published and documented in the literature regarding the effect of that resource on local library catalog departments.

There are at least seven discrete sets of evaluation skills which are likely to be required in implementing and evaluating all but the most basic automated project: (1) administrative and managerial; (2) business and financial; (3) department or function operations; (4) computer and data processing, telecommunications; (5) micrographics and publishing; (6) information science and technology; and (7) market evaluation-user needs. It would be a broad library staff which would supply the range of talents to direct a performance evaluation task. It is necessary in all but the largest libraries for the library administration to draw on a group of resources to evaluate

properly the success of a project. This does not particularly suggest the development of a committee to manage a project—generally a disastrous approach—but rather the periodic use of a number of persons representing special skills to review plans and performance. There are five general sources of such skills, each with its advantages and disadvantages.

1. *Library staff.* There is no substitute for involving staff who will be even marginally affected by the new system in the planning and evaluation (not all affected staff, but some representative from each class of staff). However, it must be recognized that staff in a particular department are not necessarily the most competent to evaluate the library's overall goals for a project.
2. *Parent or sister agencies.* Departments in the university, or agencies in public government, can often provide a most useful backup in evaluation. This is often a resource underutilized by library administrations where past experience has been, "...they don't understand library operations." While this statement is true enough historically, in the tight funding atmosphere of the 1970s the effort would seem to be politically most valuable.
3. *User groups.* Less is known about the practice and desires of library users than about any other element affecting design and function of an automated system. While formal and scientifically valid user studies are quite expensive and beyond the resources of most libraries, informal support and advice from politically well-chosen groups of users can be effective in stopping irrelevant design features. The staff is also a special user group of most automated systems, and their special problems as users should be considered in a manner separate from the authority of the system to control their activities as employees.
4. *Consultants.* A prejudiced view from the consultant's perspective is that an outside formal consultant (or one who is not affiliated with the organization and is paid under a formal contract or agreement) can save a project time and money if one or more of the following is needed: (a) education regarding information technology and the state of practice of specific applications; (b) budgeting and planning for a practical and feasible implementation and project design; (c) design expertise in information storage and retrieval, particularly in non-traditional methods; and (d) management review regarding the organization for, and implementation of, a project plan. A limiting factor is that the use of a consultant for a very short period of time generally does not allow for development of very specific guides and recommendations; from one to six man-months during the entire project period is a reasonable level of involvement for many major projects. Thus the use of consultants can be expensive; a reasonable budget might be 5-10 percent of the project effort.

5. *Vendors.* There is some tendency in the library profession to treat the vendor as an adversary whose primary goal is plundering the budget of the local library, particularly if the vendor is a commercial firm. The nature of the library profession makes this an unwarranted assumption; librarians as much as any other group of professional public agency employees have an extensive and intensive communications network. Poor or misrepresented service to a library rapidly becomes known to the library community as a whole. The vendor can be an extremely helpful resource in the planning and evaluation of a system. Control is, of course, necessary and contracts are discussed below to that end; do not, however, exclude the vendor from professional staff evaluation meetings unless the agenda is specifically a review of his performance.

### **Contracts in Performance Evaluation**

A proper contract stripped of legal and purchasing verbiage, which necessarily gets added in formal bid contracts, is simply a written specification of what is expected from the automated project. A vendor contract and an internal-project planning document are the same from an operating point of view; the sanctions for nonperformance simply differ.

The library should draw up a contract whether working with internal development, a sister data processing agency, or an outside vendor. It should include all specified products and services ("deliverables") and should include schedules and procedures of modification. The contract is the sole place where the adversary approach to performance should prevail. A vendor contract is a legal document; for agreements within the library's organization, a contract is the definition of commitments.

Do not include undefined, hoped-for goals in the contract document unless both parties to the contract clearly understand that no evaluation of performance is possible on those elements. It may be useful, particularly in bid contracts, to specify open-ended options to be specified mutually at a later time (e.g., in a book catalog contract, "and the production of such special indexes as shall be mutually agreed upon"). This technique is basically useful only for providing some extra flexibility and avoiding rebid situations for minor changes. It cannot specify price and performance on uncompleted products. On the other hand, do include in a contract specific performance benchmarks: maximum number of seconds per circulation transaction, maximum number of days turnaround for book catalog production, etc. It should, however, be recognized that sanctions specified for nonperformance are basically only protections against disaster, and not a useful operating management tool. Any benchmarks specified should only have legal status when measured in actual operating conditions.

The function of a contract (particularly a public bid) is to measure price and performance among qualified vendors, and to eliminate non-qualified responses. The bid document will not serve this function unless the library specifies clearly in the document what bases are used for evaluation of vendors. This means the library must investigate sufficiently in order to identify the truly qualified vendors before sending out the bid specification. Then the bid design must be reviewed to make sure it does not arbitrarily exclude qualified vendors and does not allow responses from nonqualified vendors.

If it is not possible to determine what is available a priori, it is the practice of some agencies (particularly large federal operations) to issue a "request for proposal," in response to which vendors can describe their services and suggest actions for the agency. This is only practicable for the largest libraries who can invest considerable staff time and funding for a major contract, because of the cost of responding to an open-ended "RFP." Most libraries are well advised to invest their own time talking personally to vendors before issuing a bid document.

There are at present three general problems with evaluating contracts for the provision of automated library services:

1. *Nothing is identical.* It is unlikely that two responses to a contract specification will be identical. We simply do not have a standard set of descriptions of library operations, much less a standard set of computer applications programs. This situation can be expected to persist in the next decade, thus providing plenty of chances for consultants and others to attempt to make something other than lemons from the apples and oranges available with which to work.
2. *Competition is sparse.* Only in the area of contract book (or microform) catalog services can one find as many as ten organizations providing nationwide services to libraries. We will perhaps see three to five on-line network systems become available within the next five years, but generally with regional emphasis. In acquisitions there are a few software or turnkey systems that are beginning to replace a previous generation of simple data collection devices.
3. *Reality is elusive.* Most libraries without special expertise will gain little in investigating standard contract products through the route of field visits or correspondence with vendors' user libraries. The problem is even more difficult if the library chooses to develop a custom product with a contractor (agency or vendor) without previous experience.

### When Disaster Looms

The best early warning system of real trouble in a system implementa-



tion is the experience that comes from having lived through one or more previous disasters. In the absence of that chastening background, some brief rules may help. Disaster may be on the way if:

1. You can't understand in English what you are buying in the computer system.
2. You can't get a firm schedule with real dollar penalties for late delivery.
3. You can't get firm estimates on processing or related operating costs.
4. Nothing comes out of any step of development for more than three months.
5. Your developer tells you library problems are really very simple!

### **Review of Program Operation**

So far we have been considering performance measures as they relate to the planning and implementation of an automated system. We have been concerned with the potential of the projected system to meet the institution's goals in terms of tasks and costs. Now we want to measure the results of all this activity—measurement which can be accomplished through review of program operation.

This review is only possible if we have gone through the previously described rigorous procedures—defining operations, determining costs, and formulating goals—prior to the implementation of the new system. Without having completed these steps, we would not only be trying to compare manual apples with automated oranges, but with hybrid strawberries as well. The purpose of this review is to give us a measure of the success or failure of the new, automated system in achieving the goals set for the program.

Success or failure can be measured in three ways which may be assigned proportionately to the desired goals: costs, performance, and benefits. The easiest measurement is that of costs, particularly if we have applied standard methodology to the determination of manual cost elements (workload, staffing, processes, resources, etc.) which may serve as benchmarks against which to measure costs of these same elements in an automated system. Similarly, actual development costs can be accurately determined and measured against projected development costs and costs of comparable systems which have been reported in the literature. This type of hard data can give management a gross estimate of the success or failure of the new system within the framework of established budgetary limitations.

The more difficult task is to obtain and, if possible, to quantify soft data regarding the success or failure of the system in achieving program goals at other than the cost level. (We are deliberately including the element of failure which is conspicuously absent in the literature, yet potentially is of even more informational value than the many records of successful, or at

least viable, operation.) These soft data fall into the categories of performance and benefit, which have been characterized by Orr<sup>7</sup> as measures of capability, utilization (service loads) and value. The first of these two are based on criteria which can indeed be measured through the answers to such questions as: What is the number of file items searched in a specified time? What is the number of outputs produced in a specified time? What is the number of files eliminated? There are many more questions which knowledgeable library staff can readily identify.

Benefit, or value, is a more difficult concept to quantify, although Orr suggests a way to do so by stating: "the value of a service must ultimately be judged in terms of the beneficial effects accruing from its use as viewed by those who sustain the costs."<sup>8</sup> While agreeing with Orr in a pragmatic sense, we feel that the difficulty in quantifying benefit lies in the fact that it is a function of the user rather than of the library or its resources. For instance, one of the projected benefits of an automated circulation system is the shortening of queuing time at the charge desk—a benefit for the user rather than the library staff member who must man the desk no matter what. In an academic environment it is possible to calculate the average faculty salary and thus assign a dollar value to the time spent by the faculty member in checking out material. But what of the student?

In order to obtain hard data in this area, we must turn to the "forced quantification of uncertainty" whereby we assign relative values to discernible benefits which are usually expressed in percentages of total cost.

What is it worth to us, the library administration, to save the patron time in checking out a book or to provide the patron with item availability information? Do these benefits account for 5 percent or 50 percent of the total cost of the automated circulation system? The answers to these and similar questions will enable us to calculate a dollar value for what are usually described as intangible benefits.

This forced quantification of uncertainty provides us with additional measure by which we may assess the success or failure of the program in achieving its specified goals, both at the time of implementation and on a continuing basis. This procedure can also provide another dimension to the traditional historical data collection by libraries which may be used in the planning of new and related programs. On a secondary level, quantification can provide a format for reporting to institution management in a manner intelligible to those who know—or care to know—little concerning the library's internal operations.

### Limits of Performance Measurement

Passing mention has been given to the various techniques employed for the performance evaluation of computer systems. Most of these—mixes,

benchmarks, etc.—are based on modeling and simulation, which in turn demand the creation of software which can consume the available funds to such an extent that none are left for the testing phase. This is probably one of the major reasons that such techniques have not been extensively employed in the evaluation of library automation efforts. Another reason is that such techniques primarily measure the performance of the equipment and not of the system as a whole, which is our primary interest. Simulation has, however, been employed in those highly structured situations where it was appropriate—as in the NELINET and OCLC studies of user terminal behavior—and where sufficient financial and technical resources were available. In this paper we have concentrated on those techniques and methodologies which are readily available to library staff without extensive technical expertise, large sums of money, or the ability to mount parallel operations. In general, these techniques and methodologies are applicable to the management of any library operation.

It should be obvious that the extent of cost and task analysis and performance measurement proposed in this paper will consume a fair proportion of the library's resources, both man-hours and dollars. What has been presented here is a model which can and will be adapted to the resources and needs of the individual institution relative to the resources required for the planning and implementation of an automated system. Analysis can become an end in itself, but we are probably beyond the point where we can afford such luxuries as full-time systems staffs devoted to this process. We should like to summarize briefly and end by paraphrasing Plato in saying that "the library which is not examined is the library not worth automating."

## REFERENCES

1. Swanson, Rowena W. "Design and Evaluation of Information Systems," *Annual Review of Information Science and Technology* 10:43-101, 1975.
2. Butler, Brett, *et al.* *Improving Public Library Access; The Los Angeles County Public Library System Access Study, Final Report*. Los Altos, Calif., R&D Consultants, 1975.
3. Learman, I. "Cost-benefit Analysis in the Design and Evaluation of Information Systems," *AFIPS Fall Joint Computer Conference, Proceedings* 39:470, 1971.
4. *Ibid.*
5. Orr, Richard H., *et al.* "Development of Methodologic Tools for Planning and Managing Library Services," *Bulletin of the Medical Library Association* 56:235-67, July 1968; and 56:380-403, Oct. 1968.
6. Lipetz, Ben-Ami. "Catalog Use in a Large Library," *Library Quarterly* 42:129-39, Jan. 1972.
7. Orr, Richard H. "Measuring the Goodness of Library Services; A General Framework for Considering Quantitative Measures," *Journal of Documentation* 29:315-32, Sept. 1973.
8. *Ibid.*, p. 318.

## ADDITIONAL REFERENCES

- Balke, K.G. "Person-centered Measurement." In *Proceedings of Computer Science and Statistics* (7th Annual Symposium on the Interface). Ames, Iowa, Iowa State University, Statistics Laboratory, 1973, pp. 86-91.
- Evans, Edward, *et al.* "Review of Criteria Used to Measure Library Effectiveness," *Bulletin of the Medical Library Association* 60:102-10, Jan. 12 1972.
- Hamburg, Morris, *et al.* "Library Objectives and Performance Measures and their Use in Decision Making," *Library Quarterly* 42:107-28, Jan. 1972.
- Holt, George A., and Stern, H.C. "Cost-benefit Evaluation of Interactive Transaction Processing Systems," *AFIPS National Computer Conference* 44:687-94, 1975.
- Janes, J.D.W. "Computer-based Decision Aids: The Evaluation Problem," *Sloan Management Review* 16:17-29, Spring 1975.
- Keen, P.G.W. "Computer-based Decision Aids: The Evaluation Problem," *Sloan Management Review* 16:17-29, Spring 1975.
- Lucas, Henry C. "Performance Evaluation and Monitoring," *Computing Surveys* 3:79-91, Sept. 1971.
- Raffel, Jeffrey A. "From Economic to Political Analysis of Library Decision Making," *College & Research Libraries* 35:412-23, Nov. 1974.